

# Violation of consistency relations and the protoinflationary transition

Massimo Giovannini <sup>1</sup>

*Department of Physics, Theory Division, CERN, 1211 Geneva 23, Switzerland*

*INFN, Section of Milan-Bicocca, 20126 Milan, Italy*

## Abstract

If we posit the validity of the consistency relations, the tensor spectral index and the relative amplitude of the scalar and tensor power spectra are both fixed by a single slow roll parameter. The physics of the protoinflationary transition can break explicitly the consistency relations causing a reduction of the inflationary curvature scale in comparison with the conventional lore. After a critical scrutiny, we argue that the inflationary curvature scale, the total number of inflationary efolds and, ultimately, the excursion of the inflaton across its Planckian boundary are all characterized by a computable theoretical error. While these considerations ease some of the tensions between the Bicep2 data and the other satellite observations, they also demand an improved understanding of the protoinflationary transition whose physical features may be assessed, in the future, through a complete analysis of the spectral properties of the B mode autocorrelations.

---

<sup>1</sup>Electronic address: massimo.giovannini@cern.ch

# 1 Introduction

Inflation must have a limited duration since it cannot extend indefinitely in the past. The lack of past geodesic completeness of a quasi-de Sitter stage of expansion suggests that the initial phase of inflation can be plausibly divided into a preinflationary phase where the background geometry decelerates<sup>2</sup> (i.e.  $\dot{a} > 0$  but  $\ddot{a} < 0$ ) followed by the protoinflationary epoch of expansion when  $\ddot{a}$  changes its sign. These periods of evolution are likely to be driven by an irrotational fluid. The quantum theory of the fluctuations of gravitating, irrotational and relativistic fluids [1] has been developed even prior to the actual formulation of the conventional inflationary paradigm and in the context of the pioneering analyses of the relativistic theory of large-scale inhomogeneities [2].

The consistency relations are essential for the determination of the inflationary curvature scale as well as for the typical value of the inflationary potential [3] at horizon crossing. It has been recently argued that the consistency relations can be violated by the protoinflationary physics [4]: a protoinflationary phase containing gravitons and fluid phonons can impact differently on the tensor to scalar ratio. In this framework, the consistency relations, even if heuristically assumed in most of the experimental analyses, are not mandatory.

Mixed initial states modify the temperature and polarization anisotropies at large scales and this idea has been scrutinized along various perspectives not only in [4] but in a number of previous discussions [5, 6, 7, 8, 9, 10, 11]. Temperature-dependent phase transitions may lead to an initial thermal state for the metric perturbations. Second-order correlation effects of the scalar and tensor fluctuations of the geometry can be used to explore the statistical properties of the initial mixed quantum state [9] (see also [12]). We shall discuss the case of spatially flat models since this is the situation suggested by the observational data. The ideas reported here can be however generalized to spatially curved background geometries.

In what follows we are going to examine the violation of the consistency relations in the light of the recent Bicep2 data [13] that are seemingly in tension with other satellite observations [14, 15]. The purpose here is not to endorse a model or a mechanism but rather to point out, through specific examples, a set of physical situations that are logically plausible. The consistency relations and their implications for the determination of the inflationary scales are discussed in section 2. In section 3 the violation of the consistency relations is examined when the protoinflationary phase contains a mixture of thermal gravitons. In section 4 we consider the more general situations where the initial state contains both thermal gravitons and thermal phonons. We finally discuss the third possibility where only the phonons are in a mixed state. In section 5 we collect the concluding remarks and draw some general lessons.

---

<sup>2</sup>The scale factor of a Friedmann-Robertson-Walker line element shall be denoted by  $a$ ; the overdot denotes a derivation with respect to the cosmic time coordinate.

## 2 Consistency relations

In the conventional lore, the inflationary curvature and energy scales can be explicitly determined provided the tensor to scalar ratio  $r_T$  is fixed, for instance, by the analysis of the B mode angular power spectra. Denoting by  $\mathcal{A}_{\mathcal{R}}$  and  $\mathcal{A}_T$  the amplitudes of the scalar and tensor fluctuations of the geometry at the conventional pivot<sup>3</sup> wavenumber  $k_p$ ,

$$P_{\mathcal{R}}(k_p) = \mathcal{A}_{\mathcal{R}}, \quad P_T(k_p) = \mathcal{A}_T, \quad (1.1)$$

the tensor to scalar ratio  $r_T$ , the tensor spectral index  $n_T$  and the slow roll parameter  $\epsilon$  obey the following chain of equalities:

$$r_T = \frac{\mathcal{A}_T}{\mathcal{A}_{\mathcal{R}}} = 16\epsilon = -8n_T, \quad (1.2)$$

where  $\epsilon = -\dot{H}/H^2$  is the slow-roll parameter. The subscript of the scalar power spectrum refers to the curvature perturbations on comoving orthogonal hypersurfaces (conventionally denoted by  $\mathcal{R}$ ); this is the gauge-invariant variable customarily employed when presenting and analyzing observational data since the first WMAP data release [14] (see also [16] for some classic references on the Bardeen formalism).

Equation (1.2) defines, in a nutshell, the physical content of the consistency relations stipulating that the slow roll parameter determines simultaneously the slope of the tensor power spectrum and the tensor to scalar ratio. From Eq. (1.2) we can determine  $\epsilon$  if the value of  $r_T$  is fixed by observations at  $k_p$ .

Equation (1.2) holds in the case of single field inflationary models which are the ones conventionally confronted with the temperature and polarization anisotropies of the Cosmic Microwave Background (CMB in what follows). The reason for focussing on single-field models is not only dictated by simplicity but also by the absence of any signal of non-gaussianity in spite of all the efforts spent so far to justify and discover large non-gaussian signals in the CMB observables. The first equality appearing in Eq. (1.2) rests on the assumption that the inflaton fluctuations are the only source of scalar inhomogeneities throughout the development of the inflationary phase. Also this second assumption seems well justified in the light of the observed value of  $r_T$ . Since  $\epsilon \ll 1$  the scalar fluctuations ascribable to a source different from the inflaton  $\varphi$  can be, at most,  $\mathcal{O}(0.1)$ .

Using Eq. (1.2) the curvature scale of inflation and the typical scale of the inflationary potential are determined as

$$\frac{H}{M_P} = \frac{\sqrt{\pi \mathcal{A}_{\mathcal{R}} r_T}}{4}, \quad \frac{W}{M_P^4} = \frac{3 r_T \mathcal{A}_{\mathcal{R}}}{128}. \quad (1.3)$$

---

<sup>3</sup>The choices for the pivot wavenumber are conventional but the common ones are  $0.002 \text{ Mpc}^{-1}$  and  $0.05 \text{ Mpc}^{-1}$ .

Finally, taking the fourth root of the second relation of Eq. (1.3) and using the definition of the number of efolds the following pair of relations can be obtained:

$$E = \left( \frac{3 r_T \mathcal{A}_{\mathcal{R}}}{128} \right)^{1/4}, \quad \left| \frac{\Delta\varphi}{\Delta N} \right| = \overline{M}_P \sqrt{\frac{r_T}{8}}, \quad (1.4)$$

where  $E$  is the typical energy scale of inflation and  $|\Delta\varphi/\Delta N|$  denotes the excursion of the inflaton field  $\varphi$  with the number of efolds  $N$ . Note that  $1/\overline{M}_P = \sqrt{8\pi}/M_P$  and  $M_P = 1.22 \times 10^{19}$  GeV; both  $M_P$  and  $\overline{M}_P$  shall be employed hereunder for convenience.

The Bicep2 experiment [13] has observed a  $B$ -mode polarization of the CMB that can be well fit by the standard  $\Lambda$ CDM scenario supplemented by tensors with  $r_T = 0.2^{+0.07}_{-0.05}$ ; the value  $r_T = 0$  is disfavored at more than  $5\sigma$ . It is possible that the actual primordial component of  $r_T$  will be slightly smaller than 0.2 even if various tests were performed on the data to eliminate systematic effects and other contaminations from galactic synchrotron and from polarized dust emissions<sup>4</sup>[17].

Using Eqs. (1.3) and (1.4) together with numerical values of  $r_T$  and  $\mathcal{A}_{\mathcal{R}}$  the various scales can be written in more explicit terms and they are

$$\left( \frac{H}{M_P} \right) = 9.70 \times 10^{-6} \left( \frac{r_T}{0.2} \right)^{1/2} \left( \frac{\mathcal{A}_{\mathcal{R}}}{2.4 \times 10^{-9}} \right)^{1/2}, \quad (1.5)$$

$$\left( \frac{W}{M_P^4} \right) = 1.12 \times 10^{-11} \left( \frac{r_T}{0.2} \right) \left( \frac{\mathcal{A}_{\mathcal{R}}}{2.4 \times 10^{-9}} \right), \quad (1.6)$$

$$\left( \frac{E}{\text{GeV}} \right) = 2.23 \times 10^{16} \left( \frac{r_T}{0.2} \right)^{1/4} \left( \frac{\mathcal{A}_{\mathcal{R}}}{2.4 \times 10^{-9}} \right)^{1/4}, \quad (1.7)$$

$$\left| \frac{\Delta\varphi}{\Delta N} \right| = 3.1 \times 10^{-2} \left( \frac{r_T}{0.2} \right)^{1/2} M_P. \quad (1.8)$$

Note, incidentally, that for  $\Delta N \simeq 5$ , Eq. (1.8) would imply  $\Delta\varphi \simeq \mathcal{O}(M_P)$ . According to some viewpoints, the excursion of the inflaton when the relevant scale exit the horizon is crucial to judge of the validity of an effective field theory approach to inflation.

The maximal number of inflationary efolds accessible to large-scale CMB measurements can be derived by demanding that the inflationary event horizon redshifted at the present epoch coincides with the Hubble radius today:

$$e^{N_{\max}} = \frac{(2\pi \Omega_{R0} \mathcal{A}_{\mathcal{R}} r_T)^{1/4}}{4} \left( \frac{M_P}{H_0} \right)^{1/2} \left( \frac{H}{H_r} \right)^{1/2-\gamma}, \quad (1.9)$$

where  $\Omega_{R0}$  is the present energy density of radiation in critical units and  $H_0^{-1}$  is the Hubble radius today. Equation (1.9) is a consequence of Eq. (1.2) since  $\epsilon$  must be expressed in terms

---

<sup>4</sup>After subtraction of some purported foregrounds the values of  $r_T$  may get closer to the Planck limits [15] and imply  $r_T = 0.16^{+0.06}_{-0.05}$ . It seems too soon to get to definite conclusions on this issue and, therefore, we shall prefer to set  $r_T = 0.2$  as a fiducial value for the tensor to scalar ratio. Slightly different values of  $r_T$  will not have any impact for the considerations discussed here.

of  $r_T$ . For the pivotal set of parameters of Eqs. (1.5)–(1.7), Eq. (1.9) becomes:

$$\begin{aligned}
N_{\max} &= 61.49 + \frac{1}{4} \ln \left( \frac{h_0^2 \Omega_{R0}}{4.15 \times 10^{-5}} \right) - \ln \left( \frac{h_0}{0.7} \right) \\
&+ \frac{1}{4} \ln \left( \frac{\mathcal{A}_{\mathcal{R}}}{2.4 \times 10^{-9}} \right) + \frac{1}{4} \ln \left( \frac{r_T}{0.2} \right) + \left( \frac{1}{2} - \gamma \right) \ln \left( \frac{H}{H_r} \right).
\end{aligned} \tag{1.10}$$

In Eqs. (1.9) and (1.10)  $\gamma$  accounts for the possibility of a delayed reheating terminating at a putative scale  $H_r$  smaller than the Hubble rate during inflation;  $\gamma$  controls the expansion rate in the intermediate phase. Since the reheating scale cannot be smaller than the one of nucleosynthesis,  $H_r$  can be as low as  $10^{-44} M_{\text{P}}$  (but not smaller) corresponding to a reheating scale occurring just prior to the formation of the light nuclei. If  $\gamma - 1/2 > 0$  (as it happens if  $\gamma = 2/3$  when the post-inflationary background is dominated by dust [18]),  $N_{\max}$  diminishes in comparison with the sudden reheating (i.e.  $H = H_r$ ) and  $N_{\max}$  can become  $\mathcal{O}(47)$ . Conversely if  $\gamma - 1/2 < 0$  (as it happens in  $\gamma = 1/3$  when the post-inflationary background is dominated by stiff sources [18, 19]),  $N_{\max}$  increases. Finally, if  $H_r = H$  (or, which is the same, if  $\gamma = 1/2$ ) there is a sudden transition between the inflationary regime and the post-inflationary epoch dominated by radiation. In spite of its dependence on  $\mathcal{A}_{\mathcal{R}}$  and  $r_T$ , the value of  $N_{\max}$  has then a theoretical error. Based on the previous considerations and on the maximal excursion of  $\gamma$  we can write

$$N_{\max} = 61.49 \pm 14.96. \tag{1.11}$$

If the total number of inflationary efolds  $N_t$  is larger than  $N_{\max}$  (i.e.  $N_t > N_{\max}$ ), then the redshifted value of the inflationary event horizon is larger than the present value of the Hubble radius.

Let us finally mention that when  $N_t \gg N_{\max}$  we expect, at least in conventional inflationary models, that any finite portion of the Universe gradually loses the memory of an initially imposed anisotropy or inhomogeneity so that the Universe attains the observed regularity regardless of the initial boundary conditions. The previous statement expresses, in practical terms, what is often dubbed cosmic no-hair conjecture (see e.g. [20] and references therein). This viewpoint has been questioned long ago by Barrow in a specific class of power-law inflationary backgrounds [21]. In general terms it is difficult to claim when the no-hair conjecture is valid. For instance in some classes of anisotropic inflationary models where the gauge fields are coupled to the inflaton, magnetic hairs can persist in spite of the number of efolds [20, 22]. In the conventional framework adopted here we can expect that the no-hair conjecture holds, as it can be established by carefully analyzing the problem within the gradient expansion [23]. This is therefore a relevant demand when assigning the protoinflationary inhomogeneities.

### 3 Violation of the consistency relations

The results of Eqs. (1.5)–(1.8) depend on the assumed consistency between the scalar and tensor modes of the geometry. This consistency can be violated even in the context of single field inflationary models by the protoinflationary dynamics. Suppose, for sake of simplicity, that during the protoinflationary background the phonons and the gravitons are not in the vacuum but in a mixed state [4, 6] (see also [8, 9, 10, 11]).

Assuming either thermal or kinetic equilibrium, the initial state is described by a density matrix with Bose-Einstein distribution i.e. for each species

$$\hat{\rho} = \sum_{\{n\}} P_{\{n\}} |\{n\}\rangle \langle \{n\}|, \quad P_{\{n\}} = \prod_{\vec{k}} \frac{\bar{n}_k^{n_k}}{(1 + \bar{n}_k)^{n_k+1}}, \quad (3.1)$$

where  $\bar{n}_k$  is the average multiplicity of each Fourier mode and  $|\{n\}\rangle = |n_{\vec{k}_1}\rangle \otimes |n_{\vec{k}_2}\rangle \otimes |n_{\vec{k}_3}\rangle \dots$ . The ellipses stand for all the occupied modes of the field. In particular the average multiplicities of gravitons and phonons are:

$$\bar{n}_k^{\text{ph}} = \frac{1}{e^{\omega^{\text{ph}}/k_T} - 1}, \quad \bar{n}_k^{\text{gr}} = \frac{1}{e^{\omega^{\text{gr}}/k_T} - 1}, \quad (3.2)$$

where, in units  $\hbar = c = k_B = 1$ ,

$$\omega^{\text{ph}} = k c_s, \quad \omega^{\text{gr}} = k, \quad k_T = T. \quad (3.3)$$

We shall posit that phonons and gravitons have the same temperature but this assumption can be dropped if only approximate (kinetic) equilibrium holds between the different species<sup>5</sup>.

The density matrix of the scalar and tensor fluctuations of the geometry is the direct product of the mixed quantum states of the phonons and of the gravitons  $\hat{\rho} = \hat{\rho}_{\text{phonons}} \otimes \hat{\rho}_{\text{gravitons}}$ ; the power spectra are obtained by tracing the product of the density matrix with the relevant field operators:

$$\text{Tr} \left[ \hat{\rho} \hat{\mathcal{R}}(\vec{x}, \tau) \hat{\mathcal{R}}(\vec{x} + \vec{r}, \tau) \right] = \int \frac{dk}{k} \mathcal{P}_{\mathcal{R}}(k, \tau) \frac{\sin kr}{kr}, \quad (3.4)$$

$$\text{Tr} \left[ \hat{\rho} \hat{h}_{ij}(\vec{x}, \tau) \hat{h}^{ij}(\vec{x} + \vec{r}, \tau) \right] = \int \frac{dk}{k} \mathcal{P}_{\text{T}}(k, \tau) \frac{\sin kr}{kr}, \quad (3.5)$$

where  $\hat{R}(\vec{x}, \tau)$  and  $\hat{h}_{ij}(\vec{x}, \tau)$  (with  $\hat{h}_i^i = 0 = \nabla_i \hat{h}_j^i$ ) denote the field operators corresponding, respectively, to the scalar and tensor modes of the geometry;  $\tau$  is the conformal time coordinate. If during the protoinflationary phase there are both thermal gravitons and thermal

---

<sup>5</sup>In practice we shall either assume that the phonons and gravitons have the same temperature (as it would follow from considerations of local thermal equilibrium) or different temperatures (as it would follow in the case of kinetic equilibrium).

phonons the final power<sup>6</sup> spectra are [4]

$$\mathcal{P}_{\mathcal{R}}(k, \tau) = \frac{8}{3M_{\text{P}}^4} \left( \frac{W}{\epsilon} \right) \left( \frac{k}{k_p} \right)^{n_s-1} (2\bar{n}_k^{\text{ph}} + 1), \quad (3.6)$$

$$\mathcal{P}_{\text{T}}(k, \tau) = \frac{128}{3} \left( \frac{W}{M_{\text{P}}^4} \right) \left( \frac{k}{k_p} \right)^{n_{\text{T}}} (2\bar{n}_k^{\text{gr}} + 1), \quad (3.7)$$

where the scalar and tensor spectral indices are, respectively,  $n_s = 1 - 6\epsilon + 2\bar{\eta}$  and  $n_{\text{T}} = -2\epsilon$ , with  $\bar{\eta} = \overline{M}_P^2 W_{,\varphi\varphi}/W$ . From the point of view of the Schrödinger description the quantum states involved in this problem are squeezed thermal states [6] (see also [24]).

Let us first consider, for simplicity, the situation where the phonons are absent (i.e.  $\bar{n}_{\text{ph}} = 0$ ) while the gravitons are in a thermal state. From the ratio of the tensor and scalar power spectra of Eqs. (3.6) and (3.7) we obtain, at the pivot scale  $k_p$  the following relation between  $\epsilon(k_p)$  and  $r_{\text{T}}(k_p)$ :

$$\epsilon(k_p) = \frac{r_{\text{T}}(k_p)}{16} \tanh \left( \frac{k_p}{2k_{\text{T}}} \right). \quad (3.8)$$

The explicit value of  $k_{\text{T}}$  at the present time depends on the maximal temperature of the protoinflationary epoch:

$$\begin{aligned} k_{\text{T}} &= T_{\text{max}} Q e^{-N_t} \left( \frac{2\Omega_{R0}}{\pi \mathcal{A} \mathcal{R} \epsilon} \right)^{1/4} \left( \frac{H_r}{H} \right)^{\gamma-1/2} \sqrt{\frac{H_0}{M_P}}, \\ T_{\text{max}} &= \left( \frac{45}{4\pi^3 g} \right)^{1/4} \sqrt{H M_P}, \end{aligned} \quad (3.9)$$

and  $Q$  is the fraction of  $T_{\text{max}}$  ascribable to the gravitons. Furthermore, the backreaction constraints during the protoinflationary stage demand that  $Q < 1$ . The value of  $k_{\text{T}}$  can be expressed in units of the present value of the Hubble rate:

$$\frac{k_{\text{T}}}{H_0} = \left( \frac{45 \Omega_{R0}}{2\pi^3 g} \right)^{1/4} Q \sqrt{\frac{M_P}{H_0}} \left( \frac{H_r}{H} \right)^{\gamma-1/2} e^{-N_t}. \quad (3.10)$$

Equation (3.10) implies an even more stringent relation between  $k_{\text{T}}$ , the total number of efolds  $N_t$  and the critical number of efolds  $N_{\text{crit}}(k)$ :

$$\frac{k}{2k_{\text{T}}} = \frac{g^{1/4}}{Q} e^{N_t - N_{\text{crit}}(k)}, \quad (3.11)$$

where  $N_{\text{crit}}(k)$  is defined as:

$$N_{\text{crit}}(k) = 66.25 - \ln \left( \frac{k}{0.002 \text{Mpc}^{-1}} \right) + \left( \gamma - \frac{1}{2} \right) \ln \left( \frac{H_r}{H} \right) + \frac{1}{4} \ln \left( \frac{h_0^2 \Omega_{R0}}{4.15 \times 10^{-5}} \right). \quad (3.12)$$

---

<sup>6</sup>The two mode functions for the two polarizations of the graviton coincide; moreover, the initial fluid phonons are the normal modes of the protoinflationary fluid that can be quantized in terms of the Lukash variable [1] which is nothing but the curvature perturbation on comoving orthogonal hypersurfaces and has been derived in the case of an irrotational fluid.

Equation (3.8) implies that  $16\epsilon(k_p) \neq r_T(k_p)$ . The consistency relations are then violated and, as it follows from Eq. (3.11),  $k_p/k_T$  is a function of  $N_t - N_{\text{crit}}(k_p)$ . The function parametrizing the relation between  $\epsilon$  and  $r_T$  can be written as  $\mathcal{F}^2(N_t - N_{\text{crit}}, g, Q)$  where

$$\mathcal{F}(x, g, Q) = \sqrt{\tanh\left(\frac{g^{1/4}}{Q}e^x\right)}. \quad (3.13)$$

Note that  $N_t$  is unknown while  $N_{\text{crit}}(k_p)$  is affected by a theoretical error comparable to the one of Eqs. (1.10) and (1.11) and coming from the different post-inflationary histories. Both numbers can be fixed by plausible theoretical guesses but they are phenomenologically undetermined. Thus, the inflationary scales previously introduced can be expressed as follows:

$$\left(\frac{H}{M_P}\right) = 9.70 \times 10^{-6} \left(\frac{r_T}{0.2}\right)^{1/2} \left(\frac{\mathcal{A}_{\mathcal{R}}}{2.4 \times 10^{-9}}\right)^{1/2} \mathcal{F}(N_t - N_{\text{crit}}, g, Q), \quad (3.14)$$

$$\left(\frac{W}{M_P^4}\right) = 1.12 \times 10^{-11} \left(\frac{r_T}{0.2}\right) \left(\frac{\mathcal{A}_{\mathcal{R}}}{2.4 \times 10^{-9}}\right) \mathcal{F}^2(N_t - N_{\text{crit}}, g, Q), \quad (3.15)$$

$$\left|\frac{\Delta\varphi}{\Delta N}\right| = 3.1 \times 10^{-2} \left(\frac{r_T}{0.2}\right)^{1/2} \mathcal{F}(N_t - N_{\text{crit}}, g, Q) M_P, \quad (3.16)$$

where  $r_T \equiv r_T(k_p)$  and  $N_{\text{crit}} \equiv N_{\text{crit}}(k_p)$ . Note that  $N_{\text{crit}}(k_p) = 66.25$  in the sudden reheating approximation (i.e.  $\gamma = 1/2$  and  $H = H_r$ ). If we assume that the reheating is not instantaneous but delayed by a long post-inflationary phase stiffer than radiation we shall have that Eqs. (3.14) and (3.15) hold with a different  $N_{\text{crit}}(k)$ . In practice we shall consider the same excursion of 15 efolds already discussed in the case of  $N_{\text{max}}$  (see Eq. (1.11)) and bear in mind that  $N_{\text{crit}}(k_p) \simeq 66.25 \pm 15$ .

The properties of  $\mathcal{F}(x, g, Q)$  depend mildly on  $g$  and  $Q$  and more crucially on  $N_t$  and  $N_{\text{crit}}$ . Let us therefore choose  $g = 2$  for the two polarizations of the graviton and  $Q = 0.1$  and let us assume the sudden reheating approximation (i.e.  $N_{\text{crit}} = 66.25$  for the fiducial set of parameters). Thus for  $N_t = 50$  we have, from Eqs. (3.14), (3.15) and (3.16) that:

$$\left(\frac{H}{M_P}\right) = 9.90 \times 10^{-9}, \quad \left(\frac{W}{M_P^4}\right) = 1.16 \times 10^{-17}, \quad \left|\frac{\Delta\varphi}{\Delta N}\right| = 3.16 \times 10^{-5} M_P. \quad (3.17)$$

Let us finally consider a far more extreme situation, namely the case where the reheating is delayed down to the nucleosynthesis scale and the expansion is stiffer than radiation (i.e. for instance  $\gamma = 1/3$  in Eq. (3.12)). Then Eq. (3.17) becomes

$$\left(\frac{H}{M_P}\right) = 5.47 \times 10^{-12}, \quad \left(\frac{W}{M_P^4}\right) = 3.57 \times 10^{-24}, \quad \left|\frac{\Delta\varphi}{\Delta N}\right| = 1.75 \times 10^{-8} M_P, \quad (3.18)$$

always for a total number of efolds  $N_t = 50$ .

Let us remark, incidentally, that an excursion of the inflaton of  $\Delta N \simeq 5$  or even 10 does not hit severely the Planckian boundary. The variation of  $Q$  corresponds to a lower



temperature of the gravitons in units of  $T_{\max}$ . If  $Q$  diminishes, for instance, by two orders of magnitude (from 0.1 to  $10^{-3}$ ) the overall effect on the curvature scale corresponds to one order of magnitude (i.e.  $10^{-8} \rightarrow 10^{-7}$  in Eq. (3.17) and  $10^{-12} \rightarrow 10^{-11}$  in Eq. (3.18)).

The total number of efolds is usually considered, for practical purposes, between 50 and 60. For instance the Planck collaboration [15] gives for the number of efolds a possible excursion between 50 and 60. A growth in the total number of efolds increases the inflationary scales; for instance when  $N_t = 60$  and in the sudden reheating approximation (i.e.  $N_{\text{crit}} = 66.25$ ) we shall have that  $H/M_P = 1.46 \times 10^{-6}$ ,  $W/M_P^4 = 2.5 \times 10^{-13}$  and  $|\Delta\varphi/\Delta N| = 4.69 \times 10^{-3}$ . These figures are a bit smaller than (but of the same order of) the ones given in the previous section.

In the limit  $N \gg N_{\text{crit}}$  the consistency relations are recovered since, as it can be explicitly checked,

$$\lim_{N_t \gg N_{\text{crit}}} \mathcal{F}(N_t - N_{\text{crit}}, g, Q) \rightarrow 1. \quad (3.19)$$

Equation (3.19) agrees with the no-hair conjecture [20, 21, 23] and it is consistent with the whole approach. Arbitrary modifications of the initial state violating the no-hair conjecture may lead to misleading conclusions unless the features of the model allow for such a violation as speculated in the past [20] and also more recently [20, 22] in different frameworks. Unfortunately neither  $N_{\text{crit}}$  nor  $N_t$  are fixed (or even bounded) by the no-hair conjecture.

The relevance of the initial conditions of large-scale fluctuations in the determination of the excursion of the scalar field have been reported in [12] in an implicit model suggesting that the effective theory can be saved if the initial state is a mixed state. We agree with the idea of [12] insofar as the initial mixed states may break the consistency relations since this is the suggestion already discussed in [4] (see also [6]). We do not agree, however, with the absence of a critical number of efolds following, on a general ground, from the validity of the no-hair conjecture in the conventional set-up of single field inflationary models. We finally disagree with the statement that the presence of a mixed state in the initial conditions is sufficient to guarantee a strong violation of the consistency relations (see, in particular, the discussion of the following section).

It would be tempting to identify  $N_{\text{crit}}$  with  $N_{\max}$  since numerically the two quantities are roughly coincident. However, at the risk of being pedantic it is appropriate to remark that since Eqs. (3.14) and (3.15) depend on the number of efolds the determination of  $N_{\max}$  is more involved than in the case when the consistency relations are valid. In particular the relation that determines  $N_{\max}$  is given, in this case, by

$$e^N \left( \frac{H}{M_P} \right)_N^{\gamma-1} = \left( 2\Omega_{R0} \right)^{1/4} \left( \frac{H_r}{M_P} \right)^{1-\gamma} \left( \frac{M_P}{H_0} \right)^{1/2}, \quad (3.20)$$

where the subscript at the left hand side reminds that  $(H/M_P)$  depends, this time, on the number of efolds. Equation (3.20) is not an algebraic equation but it can be solved in three different limits (i.e.  $N \gg N_{\text{crit}}$ ,  $N \ll N_{\text{crit}}$  and  $N = \mathcal{O}(N_{\text{crit}})$ ). Notice, finally, that when  $\gamma = 1/2$  and the consistency relations are restored Eq. (3.20) gives exactly Eq. (1.9).

## 4 Thermal phonons and thermal gravitons

The results obtained in the previous section suggest that an initially mixed state leads to a violation of the consistency relations provided the total number of efolds does not exceed the critical number of efolds. We shall now address a slightly different question and ask if the presence of a mixed state during the protoinflationary stage is also sufficient to guarantee a violation of the consistency relations.

There is an amount of fine-tuning in assuming that only the gravitons are thermal while the phonons are not. Reversing the argument we could say that the amount of breaking of the consistency relations reflects our ignorance on the total duration of the inflationary phase but also some sort of postulated asymmetry in the initial conditions of the large-scale fluctuations. While it may well be that this is exactly what the observational data demand it is nonetheless interesting to relax this assumption.

Thus, if thermal gravitons and thermal phonons are simultaneously present all the considerations developed in the previous section can be repeated with few main differences. The relation between  $\epsilon(k_p)$  and  $r_T(k_p)$  (and the consequent breaking of the consistency relations) is different from the one of Eq. (3.8) and it is given by

$$\epsilon(k_p) = \frac{r_T(k_p)}{16} \frac{\tanh\left(\frac{k_p}{2k_{Tg}}\right)}{\tanh\left(\frac{c_s k_p}{2k_{Tph}}\right)}, \quad (4.1)$$

where  $c_s$ , as already mentioned, is the sound speed of the phonons. In Eq. (4.1) we allow for different thermal wavelengths of the phonons and of the gravitons. Unlike section 3 we have that  $g \geq 3$  (since also the phonons should be counted as thermal species); finally Eq. (3.13) gets modified as follows:

$$\overline{\mathcal{F}}(x, g, Q, c_s) = \frac{\mathcal{F}(x, g, Q)}{\mathcal{F}(x, g, Q/c_s)}. \quad (4.2)$$

The form of Eq. (4.2) suggests a much less important effect on the inflationary curvature and energy scales.

More specifically, the analog of Eqs. (3.17) and (3.18) does seem to depend on  $N_{\text{crit}}$  and it is given by

$$\frac{H}{M_P} = 1.25 \times 10^{-5}, \quad \frac{W}{M_P^4} = 1.93 \times 10^{-11}, \quad \left| \frac{\Delta\varphi}{\Delta N} \right| = 4.07 \times 10^{-2}, \quad (4.3)$$

for  $N_t = 50$ ,  $g = 3$  (the two polarizations of the graviton plus the phonon) and  $c_s = 1/\sqrt{3}$  (i.e. in the case of a preinflationary phase dominated by radiation).

The rationale for the previous finding stems from the  $\overline{\mathcal{F}}(x, g, Q, c_s)$  that implies a weak breaking of the consistency relations. While for  $N_t \gg N_{\text{crit}}$  the consistency relations are recovered, when  $N_t < N_{\text{crit}}$  we can roughly approximate

$$r_T(k_p) \simeq 16 c_s \epsilon(k_p) \left( \frac{T_g}{T_{ph}} \right). \quad (4.4)$$

In thermal equilibrium  $T_g \simeq T_{ph}$  and the breaking is therefore proportional to the sound speed  $c_s$ .

As suggested in [4] it is tempting to speculate that future measurements of the tensor spectral index from the slope of the B mode polarization can decide if and how the consistency relations are broken by the initial conditions. In spite of this we must also admit that by relaxing the tuning of the initial conditions and by allowing for the presence of thermal phonons the violation of the consistency relations becomes progressively less relevant.

A complementary way of addressing this issue is through the running of the spectral index. The running of the spectral index implied by this type of models has the same qualitative features for the scalar and for the tensor modes of the geometry with the difference that the scalar spectral index is directly measured while the tensor spectral index can only be inferred from the consistency relations, if valid. In particular, defining the ratio  $\kappa = k_p/(2k_T)$  we have

$$n_s = 1 - 6\epsilon + 2\bar{\eta} + q_s + \frac{1}{2}\alpha_s \ln(k/k_p), \quad (4.5)$$

$$n_T = -2\epsilon + q_T + \frac{1}{2}\alpha_T \ln(k/k_p), \quad (4.6)$$

where, in the case discussed here, the running parameters are

$$\alpha_s = \frac{4c_s\kappa}{\sinh 2c_s\kappa} \left( -1 + \frac{2c_s\kappa}{\tanh 2c_s\kappa} \right), \quad (4.7)$$

$$\alpha_T = \frac{4\kappa}{\sinh 2\kappa} \left( -1 + \frac{2\kappa}{\tanh 2\kappa} \right), \quad (4.8)$$

while  $q_s$  and  $q_T$  are given by

$$q_s = -\frac{c_s\kappa}{\cosh c_s\kappa \sinh c_s\kappa}, \quad q_T = -\frac{\kappa}{\cosh \kappa \sinh \kappa}; \quad (4.9)$$

$q_s$  and  $q_T$  go both to 0 for  $\kappa \gg 1$  and to  $-1$  for  $\kappa \ll 1$ ;  $\alpha_s$  and  $\alpha_T$  are both positive. The running of the scalar spectral index does not seem to relax the compatibility between the Bicep2 data and the Planck upper bound on the tensor to scalar ratio, since, naively, the running is always positive (instead of negative) and, furthermore, the constant contribution changes the spectral slope too radically at large scales especially in the case of the scalar modes. It is however not clear, at the moment, if this conclusion applies also in the present case where the consistency relations are violated since the determination of the running of the spectral index from the fits assumes the validity of the consistency relations.

Let us mention, for sake of completeness that we have, in principle, a third final possibility stipulating that the breaking of the consistency relations instead of being concentrated in the graviton sector, or equally shared between phonons and gravitons is rather due to thermal phonons alone. In this third case the analog of Eqs. (3.14), (3.15) and (3.16)

$$\left( \frac{H}{M_P} \right) = 9.70 \times 10^{-6} \left( \frac{r_T}{0.2} \right)^{1/2} \left( \frac{\mathcal{A}_{\mathcal{R}}}{2.4 \times 10^{-9}} \right)^{1/2} \mathcal{F}^{-1}(N_t - N_{\text{crit}}, g, Q/c_s), \quad (4.10)$$

$$\left(\frac{W}{M_P^4}\right) = 1.12 \times 10^{-11} \left(\frac{r_T}{0.2}\right) \left(\frac{\mathcal{A}_{\mathcal{R}}}{2.4 \times 10^{-9}}\right) \mathcal{F}^{-2}(N_t - N_{\text{crit}}, g, Q/c_s), \quad (4.11)$$

$$\left|\frac{\Delta\varphi}{\Delta N}\right| = 3.1 \times 10^{-2} \left(\frac{r_T}{0.2}\right)^{1/2} \mathcal{F}^{-1}(N_t - N_{\text{crit}}, g, Q/c_s) M_P. \quad (4.12)$$

According to Eqs. (4.10), (4.11) and (4.12) the inflationary scales get larger than the conventional values. This is due to the fact that the tensor to scalar ratio instead of being determined by Eq. (3.8) is now given by

$$r_T(k_p) = 16\epsilon(k_p) \mathcal{F}^2(N_t - N_{\text{crit}}, g, Q/c_s). \quad (4.13)$$

In spite of minor differences due to the sound speed we can say that the breaking of the consistency relations goes actually in a direction that is opposite to the one to the one suggested by the observational data and, in this sense, it is purely academic.

## 5 Concluding remarks

During the protoinflationary transition the consistency relations can be violated even in the case of conventional single field models. This possibility entails various theoretical uncertainties that may interfere either constructively or destructively.

The total duration of inflation is unknown and it is customarily assigned in terms of the number of efolds, i.e. the natural logarithm of the total increase of the scale factor during inflation. According to the no-hair conjecture, when inflation lasts beyond some critical number of efolds any finite portion of the Universe gradually loses the memory of an initially imposed anisotropy or inhomogeneity. Since the preinflationary phase is likely to be dominated by radiation it is plausible that the normal modes of the geometry and of the sources will be in a mixed rather than in a pure state. The simplest possibility along this line of thinking, is the one where phonons and gravitons obey a Bose-Einstein distribution. The critical value of inflationary efolds is then determined by the temperature of the initial mixed state. When the total number of efolds greatly exceeds the critical value, according to the no-hair conjecture, the Universe attains the observed regularity regardless of the initial boundary conditions.

Three extreme physical situations can be envisaged. In the first case only the gravitons are in a mixed state. In the second case phonons and gravitons are both in a thermal state, possibly with different temperatures while in the third case only the phonons are in a mixed state. The violation of the consistency relations depends on the asymmetry of the initial data: if only the gravitons are thermal the inflationary scales can be safely lowered depending on the total number of efolds and on the critical number of efolds. Conversely the breaking of the consistency relations is moderate when both phonons and gravitons are in kinetic equilibrium. This means that the initial mixed state for the cosmological perturbations is not sufficient to guarantee a sizable violation of the consistency relations.

If the tension between the Bicep2 data and the other satellite observations will persist the potential violations of the consistency relations can offer a unique handle on the nature of the initial data, as the examples reported here suggest.

At the moment the consistency relations do not follow from any empirical evidence but just from plausible arguments that can be evaded, as the present considerations demonstrate. The future observations must then devise direct tests of the consistency relations. In particular, the forthcoming programs will be essential for the accurate determination of the tensor spectral index  $n_T$  from the slope of the B-mode power spectrum. As already suggested in [4] it is tempting to speculate that independent measurements of  $r_T$  (from the B mode amplitude) and of  $n_T$  (from the B mode slope) may offer novel diagnostics of the role played by primordial phonons and gravitons in setting the initial conditions of large-scale gravitational perturbations.

## References

- [1] V. N. Lukash, Sov. Phys. JETP **52**, 807 (1980) [Zh. Eksp. Teor. Fiz. **79**, 1601 (1980)]; V. Stokov, Astron. Rep. **51**, 431-434 (2007).
- [2] E. M. Lifshitz Zh. Eksp. Teor. Fiz. **16**, 587 (1946) [J. Phys. (USSR) **10**, 116 (1946)]; E. M. Lifshitz and I. M. Khalatnikov, Adv. Phys. **12**, 185 (1963).
- [3] S. Weinberg, *Cosmology*, (Oxford University Press, Oxford 2009).
- [4] M. Giovannini, Phys. Rev. D **88**, no. 2, 021301 (2013); Class. Quant. Grav. **29**, 155003 (2012).
- [5] I. Yu. Sokolov, Class. Quant. Grav. **9**, L61 (1992).
- [6] M. Gasperini, M. Giovannini, and G. Veneziano, Phys. Rev. **D48**, 439-443 (1993).
- [7] K. Bhattacharya, S. Mohanty and A. Nautiyal, Phys. Rev. Lett. **97**, 251301 (2006); K. Bhattacharya, S. Mohanty and R. Rangarajan, Phys. Rev. Lett. **96**, 121302 (2006).
- [8] W. Zhao, D. Baskaran and P. Coles, Phys. Lett. B **680**, 411 (2009).
- [9] M. Giovannini, Phys. Rev. D **83**, 023515 (2011); Class. Quant. Grav. **30**, 015009 (2013).
- [10] I. Agullo and L. Parker, Phys. Rev. D **83**, 063526 (2011).
- [11] S. Kundu, JCAP **1202**, 005 (2012); arXiv:1311.1575 [astro-ph.CO].
- [12] H. Collins, R. Holman and T. Vardanyan, arXiv:1403.4592.
- [13] P. A. R. Ade *et al.* [BICEP2 Collaboration], arXiv:1403.3985 [astro-ph.CO]; P. A. R. Ade *et al.* [BICEP2 Collaboration], arXiv:1403.4302 [astro-ph.CO].
- [14] D. N. Spergel *et al.*, Astrophys. J. Suppl. **148**, 175 (2003); D. N. Spergel *et al.*, *ibid.* **170**, 377 (2007); L. Page *et al.* Astrophys. J. Suppl. **170**, 335 (2007); B. Gold *et al.*, *ibid.* **192**, 15 (2011); D. Larson, *et al.*, *ibid.* **192**, 16 (2011); C. L. Bennett *et al.*, *ibid.* **192**, 17 (2011); G. Hinshaw *et al.*, *ibid.* **208** 19 (2013); C. L. Bennett *et al.*, *ibid.* **208** 20 (2013).
- [15] P. A. R. Ade *et al.* [Planck Collaboration], arXiv:1303.5082 [astro-ph.CO].
- [16] J. Bardeen, Phys. Rev. **D22**, 1882 (1980); G. V. Chibisov, V. F. Mukhanov, Mon. Not. Roy. Astron. Soc. **200**, 535 (1982); J. Bardeen, P. Steinhardt, and M. Turner, Phys. Rev. **D28**, 679 (1983); J. A. Frieman and M. S. Turner, Phys. Rev. D **30**, 265 (1984); H. Kodama, M. Sasaki, Prog. Theor. Phys. Suppl. **78**, 1-166 (1984); J.-c. Hwang, Astrophys. J. **375**, 443 (1991).

- [17] U. Fuskeland, I. K. Wehus, H. K. Eriksen and S. K. Nss, arXiv:1404.5323 [astro-ph.CO].
- [18] S. M. Leach and A. R. Liddle, Phys. Rev. D **63**, 043508 (2001).
- [19] B. Spokoiny, Phys. Lett. B **315**, 40 (1993); M. Giovannini, Phys. Rev. **D60**, 123511 (1999); Class. Quant. Grav. **16**, 2905 (1999).
- [20] M. Giovannini, Phys. Rev. D **89**, 063512 (2014).
- [21] J. D. Barrow, Phys. Lett. B **187**, 12 (1987); J. D. Barrow and O. Gron, *ibid.* **182**, 25 (1986); J. D. Barrow, Phys. Rev. D **51**, 3113 (1995); **55**, 7451 (1997).
- [22] S. Kanno, J. Soda, and M.-a. Watanabe, J. Cosmol. Astropart. Phys. **12** 024 (2010); Phys. Rev. Lett. **102**, 191302 (2009).
- [23] A. A. Starobinsky, JETP Lett. **37**, 66 (1983); R. M. Wald, Phys. Rev. D **28**, 2118 (1983).
- [24] M. S. Kim, F. A. M. de Oliveira and P. L. Knight, Phys. Rev. A **40**, 2494 (1989); P. Marian and T. A. Marian, Phys. Rev. A **47**, 4474 (1993).